

学校お便り文の高頻出語彙の縦断的研究：4年生から6年生までの名詞・サ変名詞・動詞の分析

著者	今村 桜子
雑誌名	言語資源活用ワークショップ発表論文集
巻	3
ページ	84-90
発行年	2018
URL	http://doi.org/10.15084/00001640

学校お便り文の高頻出語彙の縦断的研究 —4年生から6年生までの名詞・サ変名詞・動詞の分析—

今村桜子（横浜国立大学大学院）

Longitudinal Study of High Frequency Words in the School Letters An Analysis of Nouns and Verbs in the Elementary School News Letters from 4th to 6th grades

Eiko Imamura (Yokohama National University)

要旨

首都圏の公立小学校のお便り文(3年分 712部)からコーパスを作成し、学年ごと(4年生から6年生)の語彙の違いを分析した。本研究は、学校お便り文に用いられる語を縦断的に観察することで、外国人保護者の日本語支援に役立てることを目的とする。

KH Coderで形態素解析を行ったところ、総語数は4年 56,968語、5年 106,084語、6年 77,167語。異なり語数は4年 7,420語、5年 9,935語、6年 9,395語であった。

①頻度グラフにより少数の高頻出語と多数の低頻出語が観察される。高頻出語の学習が次年度以降の読取りに効果的であると考えられる。②品詞ごとの高頻出語を抽出し、4年生の上位100語が5、6年生の上位100語に含まれる割合を分析した結果、名詞・サ変名詞・動詞で74%から83%に上がることが分かった。③サ変名詞「卒業」は、4年生で32回、5年生で66回、6年生で104回(20位)出現する。6年生に多いが、高頻出語の学習が他学年の保護者にとっても、学校文化理解や内容スキーマ活性に役立つと示唆される。

1. はじめに

「生活のための日本語：全国調査」によると、生活する外国人に質問した言語行動のうち、「学校や園からの配布物や連絡ノートを読み、必要に応じて準備する」が「日本語でできない」割合は48.2%である(国立国語研究所 2009)。母親が外国人である場合は、お便り文を同居する夫や夫の両親、子供自身に読んでもらっているケースが多い(富谷他 2011)が、自力で処理対応できないことから、家庭内の発言権が弱かったり、自尊感情が持てなかったりするケースが報告されている(伊藤 2007)。

桑原(2017)は、生活する外国人が生活場面で手にする文書を読むために、ポイントとなる漢字・語彙・表現を精選して提示すべきだと提言している。本研究は、お便り文に用いられる語を縦断的に観察することで、優先的に学ぶべき語彙を抽出し、外国人保護者¹のお便り文書の読解支援に役立てることを目的とする。

2. 先行研究

地引(2013)は、小学校配布物 293 件を形態素解析し、解析した語が旧日本語能力試験の出題基準語彙のどの級の該当語彙であるかレベル判定し、その割合を調査した。更に

¹ 本稿での「外国人保護者」とは、日本で子どもを学校に通わせる保護者であり、日本語非母語話者の方を指す。

「名詞」から頻出語彙 150 語を抽出し、それらが小学校配布物から情報を得るために必要かどうかを、外国人保護者、日本人保護者、教師の 3 者に問う意識調査を行った。その結果、「情報を得るために必要な語彙」として 32 語を抽出した。

李(2017) は、福岡など 4 市から収集した文書を分析し「学校おたよりコーパス」を構築し、総文字数 880,869 のデータを抽出した。また、中国籍の保護者からの「中国と同じ漢字を使っている、組み合わせによって(意味が)わからなくなる」との意見から、必要度の高い複合名詞を抽出している。次に、抽出した複合名詞の理解度を調査し、「外国に存在しないか、類似したモノ・コト」「母国とは別のモノ・コトを指す場合」「漢字の多義性から生じる誤解」により、お便りの理解が困難になると分析している。さらに、語彙のみを教えるのではなく、学校文化を伝えることの重要性に言及している。

3. 研究課題

本稿の課題を以下の 3 つとした。

- ①お便り文にはどのような語が用いられるか。学年ごとの総語数と異なり語数を調査し、語彙表と、頻度グラフを作成する。
- ②どの学年にも用いられる語は何か。品詞ごとの高頻出語を縦断的に観察・分析し、網羅率を調査する。3 学年で共通して用いられる「最頻出語」を抽出する。
- ③ある学年に特有の語彙²が他学年でどのように出現するか。

4. 小学校で配布されたお便り文に用いられている語彙

4. 1 研究方法

平成 25 年度から 28 年度に首都圏の公立小学校で一人の児童に配布されたお便り文を 1121 部収集した。年度途中からの収集であった平成 25 年度分を除き、1 年分の文書が揃っている平成 26 年度から 28 年度の 3 年度分のお便りを対象とし、このうち学校と PTA から使用許諾を得た文書、合計 712 部を分析対象とした。紙のお便り文を OCR ソフト「本格読取 4」でデータ化し、KH Coder³を利用して形態素解析⁴を実施し、コーパス⁵を構築した。

また、お便り文配布の目的は、学校から家庭へ連絡事項を伝達することであり、受け取った保護者は、その内容に基づいて参加不参加の意思表示をしたり、持ち物の準備をしたりするなどの適切な行動をとることが求められる。つまり、行事名や持ち物名称の意味が理解できるだけではなく、自らのすべき対応を読み取る必要がある。そのためには「提出する」などの「サ変名詞+する」や「使う」「書く」などの動詞の習得が必要であるとの考えから、本研究では名詞のみでなく、サ変名詞と動詞も分析対象とした。

得られたデータから、品詞別に出現回数を記入した抽出語リストを出力し、それを基に、名詞、サ変名詞、動詞、動詞 B の頻度表を作成した。更に、頻度 1 位から各語までの累積語数が総語数に占める割合を計算した。

4. 2 結果と考察

学年毎の語彙数を表 1 に示す。総語数は 4 年生 56,968、5 年生 106,084、6 年生 77,167。異なり語数は 4 年生 7,420、5 年生 9,935、6 年生 9,395 であった。5 年生のお便り部数が最も多いためか、品詞ごとの総語数、異なり語数共に 5 年生の語彙数が多い(表 2)。品詞別の総

² ある学年に特有の語彙とは、その学年のみが体験するコト・モノの名称のこととする。

³ KH Coder2.0。樋口耕一氏(立命館大学産業社会学部)によって制作されたテキストマイニング用ソフトウェア。形態素解析ソフト茶筌 2.0、茶筌の辞書 IPADIC2.7、統計ソフト R3.1.0 等が同梱されている。

⁴個人情報保護の観点から「固有名詞(組織、人名、地域)」を排除したのち、記号等も排除した。

⁵本研究で構築したコーパスとは、李他(2012)に依る広義のコーパスであり、筆者の日本語教育実践の根拠とすべく構築され、使用される言語資料のことである。

語数では名詞，サ変名詞，動詞 B，動詞の順に多い。しかし，異なり語数では名詞，サ変名詞，動詞，動詞 B の順となる。動詞 B 上位の「する」「なる」「ある」「できる」等の基本語は，それぞれの出現回数が特に多いためと考えられる。KH Coder ではこれら基本語をひらがな表記の動詞「動詞 B」として別の品詞にしているため，漢字を含む動詞の出現状況が把握しやすい。

表 1 4～6 年生のお便り部数と語彙数

	4 年生	5 年生	6 年生
お便り部数	212	268	232
総語数	56,968	106,084	77,167
異なり語数	7,420	9,935	9,395

表 2 4～6 年生のお便り 品詞別出現語彙数

		4 年生	5 年生	6 年生
名詞	総語数	16507	28867	20074
	異なり語数	2422	3298	2951
サ変名詞	総語数	9189	17913	12760
	異なり語数	974	1263	1225
動詞	総語数	4410	9844	6532
	異なり語数	721	967	885
動詞 B	総語数	6319	12866	9573
	異なり語数	463	609	573

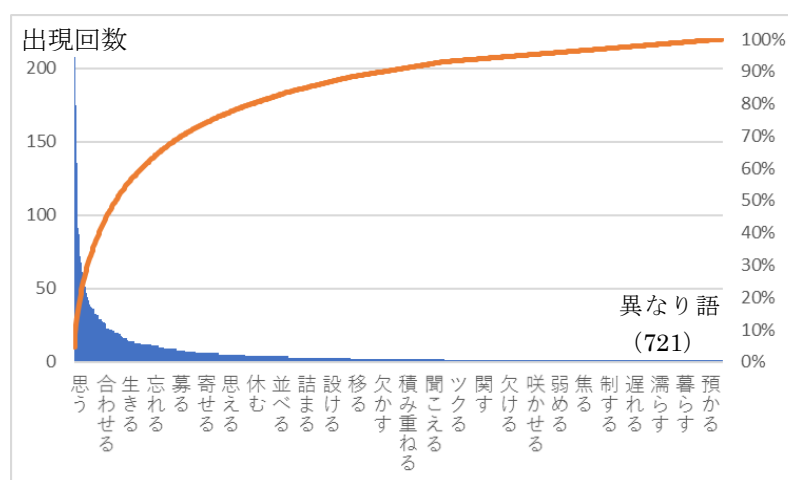


図 1 4 年生 動詞 頻度表

縦軸に出現回数，横軸に頻度をとって頻度別の棒グラフを作成すると，L字型の分布を表す(図 1)。少数の高頻度語と多くの種類の低頻度語⁶で構成されている。例えば，4 年生の動詞の総語数は 4410 で異なり語数は 721 であるが，100 位「慣れる」までで総語数に占める語の割合は 66%となり，199 位「見せる」で 80%となる。ここから，頻度の高い語を優先的に学習することで，一定の教育効果を生むと考察する。学習する外国人保護者にとっては急に理解できる語が増えたように感じられ，動機づけが高まる効果があるのではないかと。

一方で，順位の低いものを学習しても，実際に受け取るプリントに登場しない確率が増え

⁶ジップ (ZIP) の法則。日本語以外の言語調査でも確認される (萩野・田能村 2011)。

ていく。いくら学んでも上達したという実感が得られにくいという状況が生まれてしまうと予想される。しかしながら、地引(2013)で指摘される通り、頻度の低い語の中にも内容理解や行動につながる重要語が含まれる場合がある。従って、低頻度語については、お便りに出てきたときにその都度辞書を引くなり、周りの助言を得るなどして、意味を理解することが有効だと指導し、大量の語彙を覚えさせようとするのではなく、適切に辞書を使用するストラテジーや、質問をするストラテジーを、合わせて指導することが有効ではないか。

5 縦断的観察による最頻出語の抽出

5. 1 研究方法

本稿のコーパスは 3 学年分のお便りの語彙の量と内容が縦断的に観察できる点が特徴である。その点を生かし、課題 1 で得られた頻度表を基に、学年毎の異同を観察し、4 年生の高頻出語上位 100 語のうち、5 年生、6 年生でも上位 100 位に入る語がいくつあるか、品詞（名詞、サ変名詞、動詞、動詞 B⁷）ごとの網羅率を調査した（表 3、図 2）。更に 3 学年全てで 100 位以内に入っている語を「最頻出語」とし、品詞ごとに抽出した。

5. 2 結果と考察

4 年生の上位 100 語と同じ言葉は、名詞では 5 年生で 74 語、6 年生で 75 語が上位 100 位に入っている。サ変名詞は 5 年生で 83 語、6 年生で 81 語、動詞は 5 年生で 82 語、6 年生で 79 語。動詞 B については、5 年生で 72 語、6 年生で 74 語が重複していた。平均値は 77.5% である。毎年多くの語が繰り返し学校お便り文に用いられていることが明らかになった。

このような、毎年用いられる語の学習は、次年度以降に保護者がお便り文を読む際の助けになると考えられる。そのため、3 学年全てで上位 100 位に入っている語を最頻出語とし、特に優先して学習する語として提示する。支援現場ではまずその効果の可能性に言及し、動機づけを高めることが肝要であるだろう。

学習順序として、2 例提案する。①4 年生の児童をもつ外国人保護者にまず 4 年生の各品詞の上位語を提示し、次年度以降にそれ以外の高頻出語彙を提示する。②最頻出語を、割合の多い名詞、動詞の順に提示する。支援現場で接するそれぞれの保護者に応じて、適切な学習順序を組み立てれば、効率的な語彙の学習が可能になるのではないだろうか。

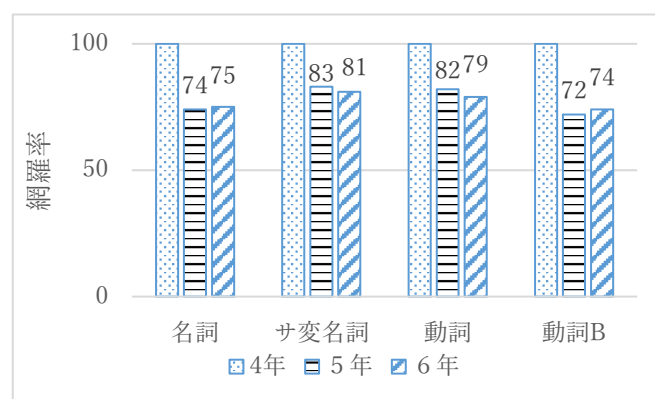


図 2 4 年生上位 100 語の 5 年生、6 年生での網羅率

品詞ごとの最頻出語と、それぞれの語の 4 年生時の出現回数を表 3-6 に示す。

⁷ 動詞 B は KH Coder 内の品詞名であり、茶笥の出力における品詞名では「動詞－自立(平仮名のみの語)」である。

表3 名詞 最頻出語

委員 533 牛乳 333 学校 317 小学校 306 先生 281 国語 254 子ども 220 算数 188
 学級 174 児童 167 学年 143 体育 143 本部 131 学期 118 クラス 111 運動会 111
 パン 109 市立 109 役員 109 大会 95 交通 93 音楽 90 食品 87 年度 85 家庭 84
 自分 80 教室 79 行事 76 地区 76 総会 75 地域 74 皆様 72 スープ 69 校外 69
 米 68 小麦粉 63 内容 61 場所 58 スポーツ 57 皆さん 56 会員 55 校長 54 情報 54
 キャベツ 53 会長 52 自転車 50 遠足 49 野菜 49 用紙 47 氏名 43 社会 40
 夏休み 39 市内 38 目標 38 コーン 37 小さじ 35 校庭 34 様子 33 チーズ 32
 体育館 31 アンケート 29 個人 29 (62語)

表4 サ変名詞 最頻出語

活動 308 給食 239 参加 195 お願い 194 指導 189 保護 179 協力 156 予定 114
 見学 112 報告 109 授業 101 お知らせ 99 連絡 96 運営 92 懇談 92 担任 91 生活 90
 理解 86 提出 84 テスト 80 教育 77 学習 74 協議 70 親睦 68 練習 68 準備 67
 成人 66 水泳 63 代表 63 注意 61 発行 60 参観 58 記入 57 会計 56 メール 55
 集金 55 開催 52 下校 51 運動 50 確認 49 当番 49 選出 47 パトロール 45 登録 42
 配布 41 お手伝い 40 広報 40 意見 37 募集 37 お話 35 試食 35 紹介 35 応援 34
 会議 34 卒業 32 電話 32 一緒 31 出席 31 担当 31 利用 31 話 31 企画 30 実施 30
 マーク 29 作成 29 体験 29 成長 28 通学 28 希望 27 開始 26 仕事 26 面談 26
 監査 25 計画 25 使用 22 (75語)

表5 動詞 最頻出語

思う 208 行う 175 食べる 136 考える 98 使う 91 見る 87 出る 72 入る 68 待つ 61
 入れる 54 作る 53 行く 52 書く 51 向ける 47 聞く 44 始まる 42 終わる 39
 出す 39 楽しむ 38 煮る 37 過ごす 36 洗う 36 感じる 33 言う 33 読む 33
 決める 32 学ぶ 29 合わせる 29 出来る 29 守る 28 加える 27 知る 27 来る 27
 含む 23 申し上げる 23 切る 23 迎える 22 取り組む 22 整える 22 深める 21
 増える 21 続く 21 伝える 20 防ぐ 20 遊ぶ 20 話す 20 頑張る 19 残す 19 違う 18
 取る 17 引き取る 16 混ぜる 16 走る 16 終わる 15 始める 14 支える 14 置く 14
 配る 14 教える 13 見守る 13 受ける 13 焼く 13 話し合う 13 開く 12 覚える 12
 願う 12 得る 12 分かる 12 歩く 12 忘れる 12 帰る 11 見える 11 乗る 11
 付ける 11 変わる 11 立つ 11 合う 10 (77語)

表6 動詞B 最頻出語

する 2085 なる 531 ある 410 できる 304 いる 244 くる 88 いう 84 いただく 71
 やく 65 つく 64 ちる 63 つける 58 こる 54 とる 54 つくる 48 つる 44 みる 44
 きる 41 よる 40 あげる 39 わかる 33 くださる 31 こめる 30 もやす 29 かむ 27
 ゆでる 26 いただける 25 かく 25 かける 23 ふる 23 やる 23 でる 22 にる 22
 たつ 18 がんばる 17 めんじる 17 おく 15 ごさる 15 いく 13 かかる 13 ける 13
 わる 13 たく 12 もつ 12 さる 11 まとめる 11 もらう 10 もる 10 ひく 9 ねる 8
 まつ 8 あう 7 しく 7 しめる 7 かう 6 (55語)

6. 一学年に特有の語の出現状況

6. 1 研究方法

課題1において得られた頻度表の名詞、サ変名詞、動詞の上位120語までを観察し、ある学年に特有の語彙として「卒業」に着目し、3学年での出現回数とその順位を比較した。

6. 2 研究結果と考察

サ変名詞「卒業」は、4年生で32回、5年生で66回、6年生で104回出現する。

「卒業」は6年生で体験する行事であることから、6年生に多いことは予想されたが、他学年でも頻度は低くはなく、上位100位以内に入っていることが注目される。「学校だより」等、全学年に向けて配布されるお便りによって、目にするものと考えられる。

ある学年に特有の言葉が他の学年にも高頻度で出現する場合があることが明らかになった。それらの語の持つ文化背景まで学ぶ機会を設けた場合には、高頻度語彙の学習が学校文化という内容スキーマを活性させることに役立つ可能性があると考えられる。

表7 「卒業」の学年毎の順位と出現回数

	4年	5年	6年
頻度順位	61位	59位	20位
出現回数	32回	66回	104回

7. おわりに

学校お便り文からコーパスを構築し、縦断的に観察・分析することで、少数の頻度上位語の数が全体に占める割合が多いことから、高頻出語の学習は効果があると考察した。また、3学年に共通して出現していた高頻出語を「最頻出語」として抽出し、優先的に学習すべき語彙として提案した。更に、一学年に特有の語が他の学年にも高頻度で出現する場合があり、これらの語の学習が内容スキーマを活性させることに役立つ可能性に言及した。

本研究は、生活する外国人の日本語支援をする立場から行った。今後は、本コーパスを教育実践に生かすため、「行事名」や「月別」などでタグ付けし、利便性を高めていくことが課題である。更に、得られた知見を基に、具体的なカリキュラムデザインや教材の作成を進めたい。

文 献

- 伊藤孝恵(2007)「国際結婚夫婦のコミュニケーションに関する問題背景：外国人妻を中心に」『言語文化と日本語 教育』33号, pp5-72
- 荻野綱男 田野村忠温(2011)『講座 IT と日本語教育 5 コーパスの作成と活用』明治書院
- 桑原陽子(2017)「初級読解教材作成を目指した非漢字系初級学習者の読解」『国立国語研究所論集』13号, pp127-141
- 地引愛(2013)「小学校配布物から情報を得るために必要な語彙の探索：使用頻度の高い語彙に注目して」『学習院大学国語国文学会誌』56号, pp76-92.
- 富谷玲子・内海由美子・仁科浩美(2012)「子育て場面で外国人保護者が直面する書き言葉の課題 ―保育園・幼稚園児の保護者を対象とした調査から―」『神奈川大学言語研究』34: pp. 53-71

- 李曉燕(2017)「外国人保護者に対する日本語支援—小学校配布プリントの特徴および「学校カルチャー語彙」の分析を通じて—」『地球社会総合科学』24, 2 号, pp1-12
- 李在鎬・石川慎一郎・砂川有里子(2012)『日本語教育のためのコーパス調査入門』くろしお出版

関連 URL

- 独立行政法人国立国語研究所日本語教育基盤情報センター学習項目グループ・評価基準グループ(2009)『生活のための日本語：全国調査』結果報告』<速報版>
http://www.ninjal.ac.jp/products/syllabus/research/pdf/seika_sokuhou.pdf